



Munich Personal RePEc Archive

# **A Conditional Value-at-Risk Based Portfolio Selection With Dynamic Tail Dependence Clustering**

De Luca, Giovanni and Zuccolotto, Paola

University of Naples Parthenope, University of Brescia

August 2013

Online at <https://mpra.ub.uni-muenchen.de/50129/>

MPRA Paper No. 50129, posted 24 Sep 2013 12:35 UTC

# **A Conditional Value-at-Risk Based Portfolio Selection With Dynamic Tail Dependence Clustering**

## **Abstract**

In this paper we propose a portfolio selection procedure specifically designed to protect investments during financial crisis periods. To this aim, we focus attention on the lower tails of the returns distributions and use a combination of statistical tools able to take into account the joint behavior of stocks in event of high losses. In detail, we propose to firstly cluster time series of stock returns on the basis of their lower tail dependence coefficients, estimated with copula functions, and secondly to use the obtained clustering solution to build an optimal minimum CVaR portfolio. In addition, the procedure is defined in a time-varying context, in order to model the possible contagion between stocks when volatility increases. This results in a dynamic portfolio selection procedure, which is shown to be able to outperform classical strategies.

**Keywords:** Copula functions, Tail dependence, Time series clustering.

## **1 Introduction**

In the multivariate analysis of financial returns, the association between extremely negative values has recently become a hot topic of research due to the recent financial crisis.

Basic portfolio theory suggests that it is not desirable to hold in portfolio assets with a strong positive association, because of the risk that a simultane-

ous drop of them could generate a considerable reduction of the value of the portfolio. However, a more refined analysis might reveal that a measure of the total association between the returns of two assets does not always provide a measure of the actual association between the lowest returns of the same assets. This implies that some popular measures of association, such as the correlation coefficient, the Kendall's tau and so on, do not always ensure the desired degree of diversification of a portfolio. On the other hand, an effective measure of the relationship between extremely low returns of two assets is the lower tail dependence because it is built taking into account the behavior of the lower tail returns, and ignoring both the central returns and the upper tail returns.

Moreover, the selection of assets with a low association between extremely low returns to compose a portfolio could be made difficult by the high number of assets available on the market.

In order to find an efficient way for selecting assets to form a portfolio keeping as low as possible the association between extreme negative returns, De Luca and Zuccolotto (2011, 2012) have proposed a multi-step procedure consisting of four steps:

- (a) the estimation of the lower tail dependence coefficients of all the possible pairs of assets;
- (b) the clustering of the assets according to the coefficients into groups characterized by high lower tail association;
- (c) the composition of portfolios selecting one asset from each cluster;

- (d) the choice among all the possible portfolios according to a specified criterion.

In this paper, in order to take into account our focus on extreme events the choice among all the possible portfolios composed of assets belonging to different clusters is made minimizing the Conditional Value-at-Risk (Rockafellar and Uryasev, 1997) rather than the variance or other dispersion measures. Moreover, we follow the idea that the lower tail dependence is not time-invariant, but has its own dynamics. As a result, the clustering of the assets is time-varying, and the possible portfolios one can compose also change over time. This approach is motivated by the recurring and erratic movements of the financial markets that can dramatically shed light on the limitations of the traditional static analyses.

The paper is organized as follows. In Section 2 the definition of lower tail dependence is given and a conveniently flexible time-varying copula model is proposed. Section 3 presents a clustering procedure based on the estimated lower tail dependence coefficients in a dynamic context, while Section 4 explains how to exploit it to compose a portfolio. In Section 5 a case study allows the comparison of two strategies of portfolio selection, named A and B. Strategy A is based on a stable clustering solution obtained summarizing all the solutions obtained at the  $T$  times of observations. Strategy B uses at each time the one-step-ahead forecast of the dissimilarity matrix to determine the clustering solution, so it is inherently dynamic. A comparison with other portfolio selection rules, such as the Markowitz mean-variance model, is discussed. Finally, Section 6 summarizes and concludes.

## 2 Dynamic tail dependence coefficients

The tail dependence is a measure of association between extremely low or high values. Let us provide some details in the bivariate case. Given two variables,  $Y_1$  and  $Y_2$  and their respective distribution functions  $F_1(Y_1)$  and  $F_2(Y_2)$ , the lower tail dependence coefficient is defined as

$$\lambda_L = \lim_{u \rightarrow 0^+} P(F_1(Y_1) < u | F_2(Y_2) < u)$$

that is the probability that  $Y_1$  assumes an extremely low value, given that an extremely low value has already occurred to  $Y_2$ .

On the other hand, if the interest is focused on very high values, the upper tail dependence coefficient is defined as

$$\lambda_U = \lim_{u \rightarrow 1^-} P(F_1(Y_1) \geq u | F_2(Y_2) \geq u)$$

that is the probability that  $Y_1$  assumes an extremely high value, given that an extremely high value has occurred to  $Y_2$ .

In a parametric approach these probabilities are model-dependent, that is the choice of a model for the bivariate set of data implies or neglects non-null tail dependence coefficients. A bivariate Gaussian model is a typical example of model which does not admit any tail dependence. Therefore, this parametric hypothesis is valid only if the assumption of no association between extreme values is reasonable.

In the analysis of bivariate financial returns, the inadequacy of the Gaussian paradigm is widely acknowledged. The tails of the empirical univariate distributions of the data are fatter than the Gaussian model, and the tails of the

empirical bivariate distributions show an association among extreme values not encountered in the bivariate Gaussian distribution.

A more complex but flexible way of describing a set of bivariate data is the use of a copula function. A bivariate copula function is defined as a function  $C : [0, 1]^2 \rightarrow [0, 1]$  such that the joint distribution function  $F(y_1, y_2)$  can be written as

$$F(y_1, y_2) = C(F_1(y_1), F_2(y_2))$$

for all  $y_1, y_2$  (see Nelsen, 2006).

The flexibility of the copula function  $C$  allows us to model the joint density separating the marginal distributions from the dependence structure.

In this case, it is easy to show that the tail dependence coefficients can be written in terms of the copula function. In particular, the lower tail dependence coefficient is given by

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}$$

while the upper tail dependence coefficient is

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u}.$$

Given the wide variety of copula functions, one can select a function with no tail dependence, or a function with only one tail dependence (lower or upper) or a function which admits both tail dependence in a symmetric or asymmetric way. In the analysis of financial returns, an association between extreme values is usually detected in both the tails. So, a natural choice is a fairly general copula function allowing a non symmetric tail dependence in the two tails. The

Joe-Clayton copula function (Joe, 1997) meets this basic requirement and is easily estimable. In the bivariate case, the Joe-Clayton copula function is given by

$$C(u_1, u_2) = 1 - \{1 - [(1 - (1 - u_1)^\kappa)^{-\theta} + (1 - (1 - u_2)^\kappa)^{-\theta} - 1]^{-1/\theta}\}^{1/\kappa}, \quad (1)$$

where  $u_i$  represents the distribution function of the  $i$ -th variable. The Joe-Clayton copula depends on two parameters,  $\theta > 0$  and  $\kappa \geq 1$ . The lower and upper tail dependence coefficients are determined by, respectively,  $\theta$  and  $\kappa$ , that is

$$\lambda_L = 2^{-\frac{1}{\theta}}$$

and

$$\lambda_U = 2 - 2^{\frac{1}{\kappa}}.$$

A time-invariant copula involves constant tail dependence. However, financial markets constitute a dynamic context exposed to many different stresses. As a result the constancy of tail dependence could be a restrictive assumption. In this work we relax this hypothesis, proposing a time-varying Joe-Clayton copula function. In particular, as our interest lies in the lower tail, that is in the relationship between extremely negative returns, we propose a time-varying model only for the parameter  $\theta$ , driving the lower tail dependence coefficient, while  $\kappa$  is kept constant. Moreover, we assume that the dynamics of  $\theta$  is affected by the past volatility of the market. In particular, denoted  $\theta_t$  and  $\sigma_t$ , respectively, the parameter and the market volatility at time  $t$ , we assume that

the time-varying Joe-Clayton copula is given by

$$C(u_{1t}, u_{2t}) = 1 - \{1 - [(1 - (1 - u_{1t})^\kappa)^{-\theta_t} + (1 - (1 - u_{2t})^\kappa)^{-\theta_t} - 1]^{-1/\theta_t}\}^{1/\kappa}, \quad (2)$$

and

$$\Delta\theta_t = \alpha (\sigma_{t-1} - \gamma). \quad (3)$$

The interpretation of  $\gamma$  is surely interesting. It can be seen as a threshold. When  $\alpha > 0$ , if the volatility at time  $t-1$  is over the threshold, then an increase of  $\theta$  is expected, that is  $\Delta\theta_t > 0$ . Viceversa, if  $\sigma_{t-1}$  is under the threshold, then  $\Delta\theta_t < 0$ . When the parameter  $\alpha$  is negative, an opposite mechanism works.

Equation (3) can be written as

$$\Delta\theta_t = \omega + \alpha\sigma_{t-1}$$

where  $\omega = -\alpha\gamma$  and finally

$$\theta_t = \omega + \theta_{t-1} + \alpha\sigma_{t-1}.$$

Then, the time-varying lower tail dependence coefficient is obtained as

$$\lambda_{Lt} = 2^{-1/\theta_t}. \quad (4)$$

### 3 Dynamic time-series clustering

#### 3.1 The clustering procedure

In this paper we refer to the clustering procedure proposed in De Luca and Zuccolotto (2011), where time series of financial returns are clustered using a



dissimilarity measure defined as

$$\delta(\{y_{it}\}, \{y_{jt}\}) = \delta_{ij} = -\log(\hat{\lambda}_L), \quad (5)$$

where  $\{y_{it}\}_{t=1,\dots,T}$  and  $\{y_{jt}\}_{t=1,\dots,T}$  denote the time series of returns of two assets  $i$  and  $j$ , and  $\hat{\lambda}_L$  is their estimated tail dependence coefficient.

Given the time series of the returns of  $p$  assets, the clustering procedure is composed of two steps. In step 1, starting from the dissimilarity matrix  $\Delta = (\delta_{ij})_{i,j=1,\dots,p}$ , an *optimal* representation of the  $p$  time series  $\{y_{1t}\}, \dots, \{y_{pt}\}$  as  $p$  points  $\mathbf{y}_1, \dots, \mathbf{y}_p$  in  $\mathbb{R}^q$  is found by means of Multidimensional Scaling (MDS). The term *optimal* means that the Euclidean distance matrix  $D = (d_{ij})_{i,j=1,\dots,p}$ , with  $d_{ij} = \|\mathbf{y}_i - \mathbf{y}_j\|$ , of the  $p$  points  $\mathbf{y}_1, \dots, \mathbf{y}_p$  in  $\mathbb{R}^q$  has to fit as closely as possible the dissimilarity matrix  $\Delta$ . The extent to which the interpoint distances  $d_{ij}$  “match” the dissimilarities  $\delta_{ij}$  is measured by an index called *stress*, which should be as low as possible. The algorithm of MDS works for a given value of the dimension  $q$ , which has to be given in input. So, it is proposed to start with the dimension  $q = 2$  and then to repeat the analysis by increasing  $q$  until the minimum stress of the corresponding optimal configuration is lower than a given threshold  $\bar{s}$ . In step 2, the  $p$  points  $\mathbf{y}_1, \dots, \mathbf{y}_p$  in  $\mathbb{R}^q$  are clustered using a  $k$ -means algorithm.

The clusters obtained with this procedure are composed of assets characterized by high tail dependence in the lower tail. De Luca and Zuccolotto (2011, 2012) show that the clustering solution obtained with this procedure can be effectively exploited for portfolio selection. The basic idea consists of avoiding to invest on assets belonging to the same cluster, in order to counterbalance

possible extreme losses. So, we should select assets by imposing the restriction that each asset belongs to a different cluster. This protects the investments from parallel extreme losses during crisis periods, because the clustering solution is characterized by a moderate lower tail dependence between clusters.

### 3.2 The dynamic clustering procedure

Through the copula function described in Section 2 we obtain time-varying estimates of the tail dependence coefficients. Given two time series of financial returns  $\{y_{it}\}_{t=1,\dots,T}$  and  $\{y_{jt}\}_{t=1,\dots,T}$ , let  $\hat{\lambda}_{Lt}$  be their lower tail dependence estimate at time  $t$ . The dissimilarity measure (5) between the two series can then be computed for each time  $t$  as

$$\delta_{ij,t} = -\log(\hat{\lambda}_{Lt}). \quad (6)$$

On the whole, given the time series of returns of  $p$  assets, we get time-varying  $(p \times p)$  dissimilarity matrices

$$\Delta_t = (\delta_{ij,t})_{i,j=1,\dots,p;t=1,\dots,T}. \quad (7)$$

So, we can sequentially apply the clustering procedure described in Section 3.1 to the matrices  $\Delta_t$ ,  $t = 1, \dots, T$ , thus obtaining a different clustering solution at each time  $t$ . The groups composition is dynamically adapted to the variations due to the changes in the copula function parameter. The dynamics of the clustering solutions can be inspected in different ways: the overall discordance between the clustering at time  $t - 1$  and that at time  $t$  can be measured by using a normalized dissimilarity index  $d_t$ , such as for example the Rand index

(Rand, 1971). Alternatively, we could examine the patterns of some pairs of assets (same cluster/different cluster) we are interested to.

Finally, the dynamic of the whole period can be summarized by computing, for each pair of assets  $ij$ , the index

$$b_{ij} = 1 - \frac{\sum_{t=1}^T I_{ij,t}}{T} \quad (8)$$

where  $I_{ij,t}$  is the indicator function which equals 1 if stock  $i$  and stock  $j$  are assigned to the same cluster at time  $t$  and 0 otherwise. The index  $b_{ij}$  denotes the fraction of the total clustering solutions with stocks  $i$  and  $j$  belonging to different clusters, a generalization of the widely used Simple Matching distance, due to Sokal and Michener (1958). So, we can perform a clustering algorithm using the distance matrix  $B = (b_{ij})_{i,j=1,\dots,p}$ , in order to summarize the  $T$  dynamic cluster solutions. The final clustering solution obtained using the distance matrix  $B$  will be called *Overall Dynamic Clustering* (ODC).

## 4 Portfolio selection

In our idea, the main challenge of clustering is the possibility to use it for building a portfolio able to protect investments during periods when extreme losses could occur simultaneously for many stocks, due to contagion phenomenons.

As pointed out in the Introduction, in this paper we improve the method proposed in De Luca and Zuccolotto (2011, 2012) firstly by using a portfolio selection procedure focused on extreme events, coherently with the tail dependence approach, and secondly by relaxing the assumption of constant tail dependence

over time.

About the first point, we propose to build portfolios by optimizing Conditional-Value-at-Risk (CVaR), a measure of risk defined by Rockafellar and Uryasev (2000) as the expected loss exceeding Value-at Risk (VaR) and better fitting in our context, where the focus is on the tails of the probability distributions. In the literature, CVaR is also called Mean Excess Loss, Expected Shortfall, or Tail VaR. Let  $\mathbf{Y} = (Y_1, \dots, Y_p)'$  be a multiple random variable with probability density  $f(\mathbf{Y})$ , describing the returns of  $p$  assets at a given time  $t$ , and  $\mathbf{w} = (w_1, \dots, w_p)'$ ,  $w_1 + \dots + w_p = 1$ , a vector of weights of the  $p$  assets in a portfolio  $P(\mathbf{Y}, \mathbf{w})$ . The loss associated to  $P(\mathbf{Y}, \mathbf{w})$  is given by  $L(\mathbf{Y}, \mathbf{w}) = -\mathbf{w}'\mathbf{Y}$ . For a given probability level  $\beta \in (0, 1)$ ,  $\text{VaR}_\beta$  and  $\text{CVaR}_\beta$  are defined respectively as

$$\begin{aligned}\text{VaR}_\beta &= \min\{\alpha \in \mathbb{R} : \Psi(\mathbf{w}, \alpha) \geq \beta\} \\ \text{CVaR}_\beta &= \frac{\int_{L(\mathbf{Y}, \mathbf{w}) \geq \text{VaR}_\beta} L(\mathbf{Y}, \mathbf{w}) f(\mathbf{Y}) d\mathbf{Y}}{(1 - \beta)}\end{aligned}$$

where

$$\Psi(\mathbf{w}, \alpha) = \int_{L(\mathbf{Y}, \mathbf{w}) \leq \alpha} f(\mathbf{Y}) d\mathbf{Y}.$$

In order to identify the set of weights  $\mathbf{w}$  minimizing  $\text{CVaR}_\beta$  for a given  $\beta$ , Rockafellar and Uryasev (2000) define the following objective function

$$S(\mathbf{w}, \text{VaR}_\beta) = \text{VaR}_\beta + \frac{\int_{\mathbf{Y} \in \mathbb{R}^p} [-\mathbf{w}'\mathbf{Y} - \text{VaR}_\beta]^+ f(\mathbf{Y}) d\mathbf{Y}}{(1 - \beta)}$$

where  $[a]^+ = a$  when  $a > 0$  and  $[a]^+ = 0$  otherwise. Given the time series of the returns of  $p$  assets  $\{y_{1t}\}, \dots, \{y_{pt}\}$ ,  $t = 1, \dots, T$ , sampled from  $f(\mathbf{Y})$ , let

$\mathbf{y}_t = (y_{1t}, \dots, y_{pt})'$  be the vector of  $p$  observations at time  $t$ . The objective function  $S(\mathbf{w}, \text{VaR}_\beta)$  can then be approximated using data as

$$\hat{S}(\mathbf{w}, \text{VaR}_\beta) = \text{VaR}_\beta + \frac{\sum_{t=1}^T u_t}{T(1 - \beta)} \quad (9)$$

where  $u_t = \max\{-\mathbf{w}'\mathbf{y}_t - \text{VaR}_\beta; 0\}$ . We have to compute the values of  $\mathbf{w}$  and  $\text{VaR}_\beta$  that minimize function (9), subject to the following linear constraints:

- $w_1 + \dots + w_p = 1$ ;
- $\mathbf{w}'\bar{\mathbf{y}} \geq y_{\min}$ , where  $\bar{\mathbf{y}} = T^{-1} \sum_{t=1}^T \mathbf{y}_t$  and  $y_{\min}$  is the minimum allowed expected return for the portfolio  $P(\mathbf{Y}, \mathbf{w})$ ;
- $u_t \geq 0$  and  $\mathbf{w}'\mathbf{y}_t + \text{VaR}_\beta + u_t \geq 0$  for  $t = 1, \dots, T$ .

The problem can be solved with linear programming. Further details about portfolio optimization with minimum CVaR objective can be found in Rockafellar and Uryasev (2000) and Krokmal *et al.* (2002).

So, we propose the following two strategies for clustering based portfolio selection:

1. Use time series of price returns of  $p$  stocks at time  $1, \dots, T$ ;
2. fit each time series with a proper univariate model accounting for possible autocorrelation and heteroskedasticity;
3. using the *i.i.d.* residuals of the marginal models, estimate the parameters of all the bivariate time-varying copula functions (a total of  $0.5p(p-1)$  bivariate distributions is estimated at this step);

4. compute the  $T$  dissimilarity matrices  $\Delta_t = (\delta_{ij,t})_{i,j=1,\dots,p;t=1,\dots,T}$  according to (7).

#### STRATEGY A

- (5A) Determine the ODC clustering solution described above, using the distance matrix  $B = (b_{ij})_{i,j=1,\dots,p}$ , with  $b_{ij}$  computed by (8). Let  $k$  be the number of clusters;
- (6A) for a fixed  $\beta$ , build all the possible Minimum CVaR Portfolios composed by  $k$  stocks not belonging to the same cluster of the ODC clustering solution;
- (7A) select the Minimum CVaR Portfolio with the lowest CVaR value.

#### STRATEGY B

- (5B) At time  $T$ , compute the one-step-ahead forecast of the dissimilarity matrix  $\Delta_{T+1}$  and determine the clustering solution for time  $T + 1$  using the two-step procedure described in section 3.1. Let  $k$  be the number of clusters;
- (6B) for a fixed  $\beta$ , build all the possible Minimum CVaR Portfolios composed by  $k$  stocks not belonging to the same cluster of the clustering solution determined in the previous step;
- (7B) select the Minimum CVaR portfolio with the lowest CVaR value.

The fundamental difference between strategy A and B lies in the nature of the clustering solution employed to determine the composition of the candidate portfolios. Strategy A uses the ODC clustering solution, built by summarizing

all the clustering solutions obtained at the  $T$  times  $1, \dots, T$ , while strategy B uses the specific instantaneous clustering solution determined at time  $T + 1$ . Thus, strategy A is based on a stable clustering solution and the corresponding portfolio does not need to be frequently updated, while strategy B is based on clustering solutions that can potentially change every day, and frequent revision is then recommended. The choice between the two strategies depends on the state of the market: strategy B can be useful during periods with high instability.

## 5 Case study

In this case study we analyse the time series of the daily price returns of the  $p = 24$  Italian stocks which have been included in FTSE MIB index during the whole period from January 3, 2006 to October 31, 2011 ( $T = 1481$ ). We firstly estimate the marginal models of the  $p$  time series by fitting data with univariate Student- $t$  AR-GARCH models, in order to take into account the possible presence of autocorrelation and heteroskedasticity. Then the copula function is estimated using the distribution functions computed on the *i.i.d.* residuals of the marginal models.

### 5.1 Estimation of dynamic tail dependence coefficients

For each pair of standardized residuals, the time-varying Joe-Clayton copula function (2) has been estimated using equation (3) for modelling  $\theta_t$ .

The copula density has been maximized using a routine written in Gauss code to obtain maximum likelihood estimates.

The estimates for the 276 pairs of residuals can be summarized as follows:

- (a) the estimates of  $\omega$  range from -0.0027 to approximately 0; in 235 cases out of 276 the estimate is negative;
- (b) the estimates of  $\alpha$  range from -0.0756 to 0.2125; in 234 cases out of 276 the estimate is positive;
- (c) the estimates of  $\kappa$  range from 1.0387 to 1.8898.

We have then computed the time-varying lower tail dependence coefficients according to (4). Figure 1 gives a rough summary of the dynamics of the 234 estimated coefficients characterized by a positive  $\alpha$ , while Figure 2 contains the dynamics of the remaining 42 coefficients with a negative  $\alpha$ .

The analyzed period is characterized by low volatility in the first half, then, at the end of 2008 a peak in the volatility is observed, followed by an uncertain period with a swing of volatility. This is reflected in the dynamics of the (lower) tail dependence. For the pairs with an estimated positive  $\alpha$  (the majority), the tail dependence is weaker in the first half, grows up rapidly at the end of the 2008, then its increase is slow until the end of the period (October 2011). In a very few cases, and only for the last part of the period of observation, the tail dependence coefficient is greater than 0.60. In general, the range of the coefficients is wide. At the beginning of the sample period (January 2006) the coefficients approximately range from 0.05 to 0.55, before the crisis from 0.01 to 0.50, at the end (October 2011) from 0.05 to 0.70. For the remaining pairs with an estimated negative  $\alpha$ , the tail dependence is approximately constant or



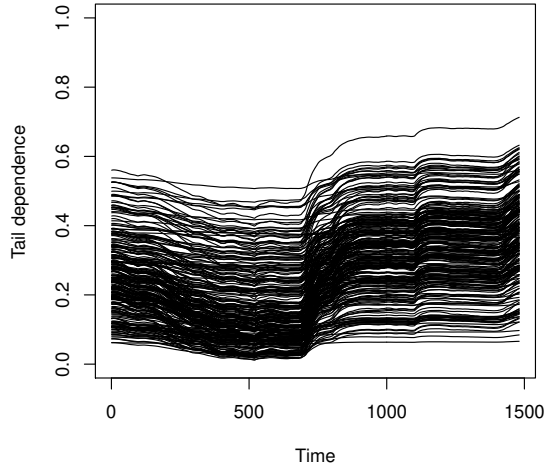


Figure 1: Dynamics of the lower tail dependence of the 234 pairs with positive  $\alpha$ .

presents very small changes also in correspondence of the peaks of the volatility. The coefficients range from 0.05 to 0.50.

Moreover, we report two cases characterized by a different sign of  $\alpha$ . The former is the pair ATLANTIA-AUTOGRILL presenting a positive  $\alpha$ , the latter the pair MPS-LUXOTTICA characterized by a negative sign of  $\alpha$ . Figure 3 depicts the estimated time-varying lower tail dependence coefficients of the two pairs, ATLANTIA-AUTOGRILL (top) and MPS-LUXOTTICA (bottom), with the estimated one-lagged volatility of the market superimposed and the estimated thresholds  $\hat{\gamma}$ . We can observe that when the volatility is below the threshold the tail dependence coefficient of the pair ATLANTIA-AUTOGRILL

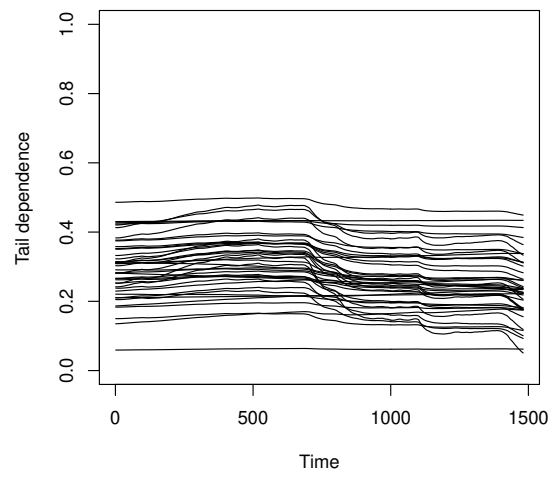


Figure 2: Dynamics of the lower tail dependence of the 42 pairs with negative  $\alpha$ .

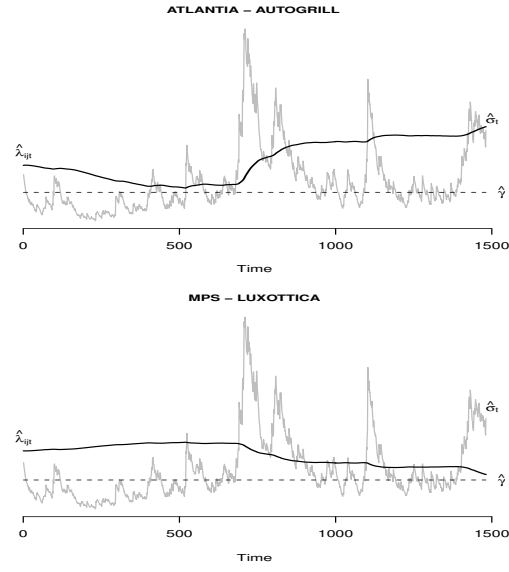


Figure 3: Lower tail dependence coefficient at time  $t$  and market volatility at time  $t-1$  for the pairs ATLANTIA-AUTOGRILL (top) and MPS-LUXOTTICA (bottom).

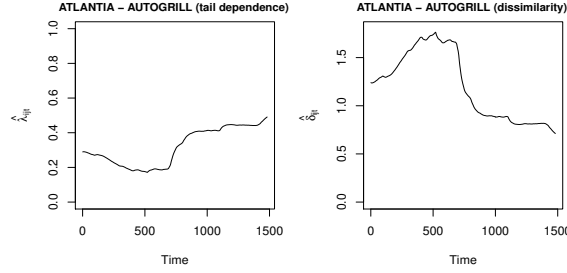


Figure 4: Lower tail dependence coefficient (left) and dissimilarity index (right) for the pair ATLANTIA-AUTOGRILL.

tends to decrease, on the other hand, when the volatility exceeds the threshold the same coefficient tends to increase. In particular, in correspondence of the middle highest peak of the volatility and, in the second part of the sample period, of the other two peaks, the coefficient shows more pronounced increases. For the pair MPS-LUXOTTICA, the opposite behavior is observed.

Finally, Figure 4 depicts the lower tail dependence coefficient (left) and the dissimilarity index (right) of the pair ATLANTIA-AUTOGRILL. In Figure 5 the same graphs are reported for the pair MPS-LUXOTTICA.

## 5.2 Dynamic clustering

After the estimation of the  $0.5p(p-1) = 276$  bivariate copula functions, the sequence of the estimated time-varying  $(p \times p)$  dissimilarity matrices  $\Delta_t = (\delta_{ij,t})_{i,j=1,\dots,p;t=1,\dots,T}$  is computed as in (7). By sequentially applying the clustering procedure described in Section 3.1 to the matrices  $\Delta_t$ ,  $t = 1, \dots, T$ , we obtain a different clustering solution at each time  $t$ .

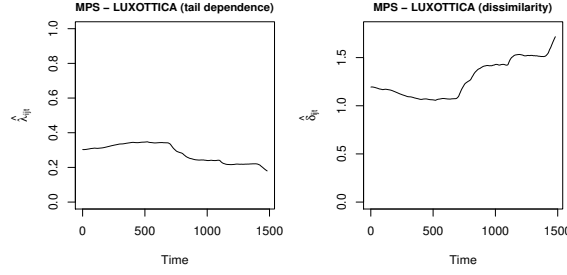


Figure 5: Lower tail dependence coefficient (left) and dissimilarity index (right) for the pair MPS-LUXOTTICA.

Computations are carried out using the R functions `isoMDS` for MDS, `kmeans` and `cclust` for  $k$ -means clustering.

From an operative point of view, since the  $k$ -means clustering solutions can be different from one run to another of the algorithm, at each time  $t$  the following procedure is used for each value of  $k$ ,  $k = 1, \dots, K$ , where  $K$  is fixed by the researcher and denotes the maximum reasonable number of clusters. In this case study we have fixed the value  $K = 10$ .

Firstly `kmeans` is executed 8 times with the Hartigan-Wong algorithm (Hartigan and Wong, 1979).

Then, the centroids of the solution with the highest ratio of the variance between over the total variance are used as optimal initial values for `cclust` which, in its turn, is executed 3 times and the best solution is chosen according to the same criterion of the highest ratio of the variance between over the total variance.

Once obtained a stable clustering solution for each value of  $k$ , the optimal

number of clusters is selected as the most voted by the following indices (for a comprehensive review, see Dimitriadou *et al.*, 2002): RL (Ratkowsky and Lance, 1978), SS (Scott and Symons, 1971), FR (Friedman and Rubin, 1967), DB (Davies and Bouldin, 1979), SSI (Dolničar *et al.*, 1999).

For each pair of stocks  $ij$ , at each time  $t$  we record the value  $I_{ij,t} = 1$  if the two stock belong to the same cluster and  $I_{ij,t} = 0$  otherwise. Figure 6 shows the scatterplots of  $I_{ij,t}$  against  $\sigma_{t-1}$  for  $i=\text{ATLANTIA}$  and  $j$  all the other 23 stocks (solid lines denote a kernel smoothing showing the basic average patterns).

For low levels of volatility the pattern is sometimes wavering, but when volatility increases, the pairs tend to be more stably assigned either to the same or to a different cluster, as a result of the stronger tail dependence characterizing some pairs of stocks during high volatility periods. Similar graphs are obtained for the other pairs of stocks. Finally, the ODC clustering solution is determined by applying a hierarchical algorithm (Ward linkage) to the distance matrix  $B = (b_{ij})_{i,j=1,\dots,p}$  with  $b_{ij}$  computed by (8). The obtained cluster dendrogram is displayed in Figure 7, with  $k = 5$  clusters selected.

### 5.3 Portfolio selection

In this section we show the results of portfolio selection according to the two strategies described in Section 3.2. We have fixed the value  $\beta = 0.05$ .

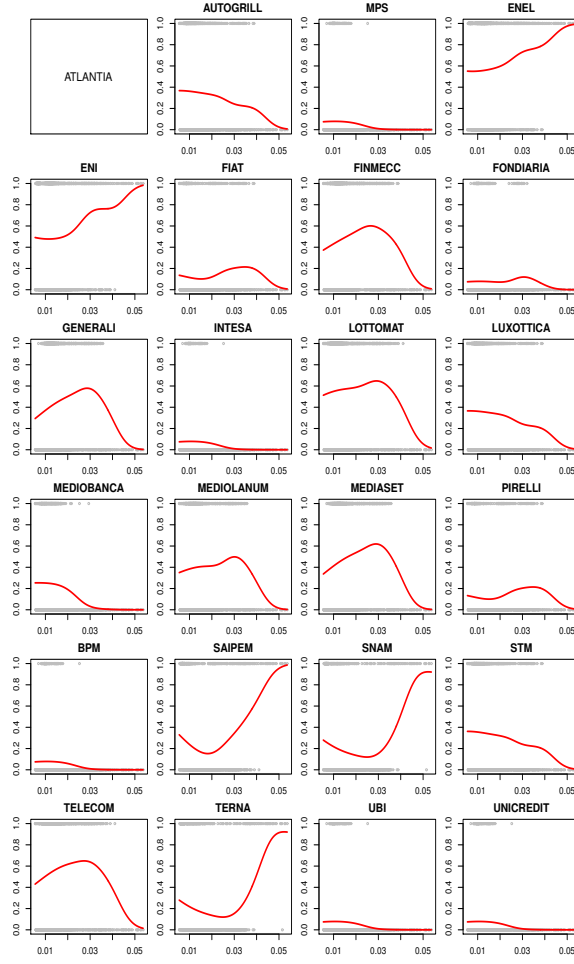


Figure 6:  $I_{ij,t} = 1$  versus  $\sigma_t$  for  $i=\text{ATLANTIA}$  and  $j$  all the other 23 stocks.

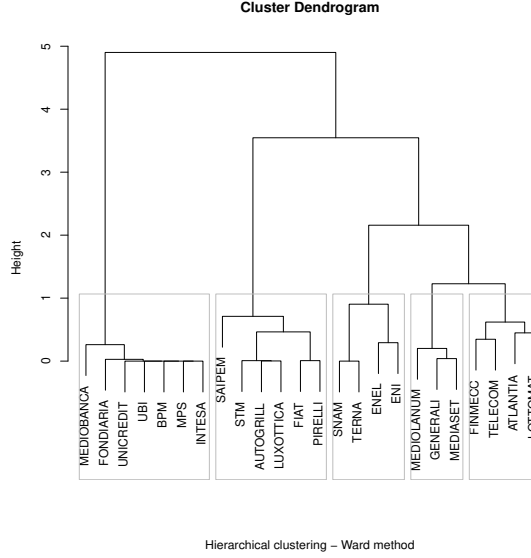


Figure 7: ODC clustering solution

### 5.3.1 Strategy A

Starting from the ODC clustering solution of Figure 7 with  $n_1 = 7$ ,  $n_2 = 6$ ,  $n_3 = 4$ ,  $n_4 = 3$ ,  $n_5 = 4$ , where  $n_i$  is the number of stocks belonging to cluster  $i$ , we can build a total number of  $n_1 n_2 n_3 n_4 n_5 = 2016$  possible portfolios of  $k = 5$  stocks using the criterion of selecting stocks not belonging to the same cluster. For each of these 2016 possible selections, we estimate the weights of the Minimum CVaR Portfolio using returns at times  $1, \dots, T$  and we finally select the solution which exhibits the lowest CVaR value (Table 1). Computations are carried out by solving the linear programming problem described in Section 3.2, using the R package `Rglpk`.

For sake of comparison, we also build the Markowitz minimum variance



Table 1: Minimum CVaR Portfolio based on ODC

<b>Stock</b>	Mediobanca	Luxottica	Snam	Generali	Atlantia
<b>weights</b>	0.0796	0.1634	0.6759	0	0.0811

Portfolio and the Minimum CVaR Portfolio using the returns at times  $1, \dots, T$  of all the  $p = 24$  stocks (Tables 2 and 3). Computations for Markowitz portfolios have been carried out using the R package `fPortfolio`.

The three portfolios (Markowitz minimum variance Portfolio, Minimum CVaR Portfolio, Minimum CVaR Portfolio based on ODC) exhibit a quite similar structure, the one based on ODC being the most parsimonious, as it requires investment on 4 stocks, while the other two results composed by 8 and 6 stocks, respectively. A barplot of the weights of the stocks in the three portfolios is displayed in Figure 8.

Finally we check the performance of the three portfolios with an out-of-sample perspective, in the period from November 1, 2011 to December 20, 2011, that is at times  $T + 1, \dots, T + 36$ . The Minimum CVaR Portfolio based on ODC tends to outperform the other two (Figure 9).

### 5.3.2 Strategy B

As pointed out in Section 3.2, portfolios built according to strategy B usually need to be frequently rebalanced, as they rely on instantaneous clustering solutions. In this case study we decide to update the portfolio composition every 5 days. So, for the out-of-sample period from November 1, 2011 to December 20, 2011, the procedure is carried out at times  $T$  (October 31),  $T + 5$  (November

Table 2: Markowitz minimum variance Portfolio

<b>Stock</b>	Atlantia	Autogrill	Mps	Enel
<b>weights</b>	0.0454	0.0384	0	0
<b>Stock</b>	Eni	Fiat	Finmecc	Fondiaria
<b>weights</b>	0	0	0.0396	0
<b>Stock</b>	Generali	Intesa	Lottomat	Luxottica
<b>weights</b>	0	0	0.0363	0.0138
<b>Stock</b>	Mediobanca	Mediolanum	Mediaset	Pirelli
<b>weights</b>	0.0790	0	0	0
<b>Stock</b>	Bpm	Saipem	Snam	Stm
<b>weights</b>	0	0	0.4448	0
<b>Stock</b>	Telecom	Terna	Ubi	Unicredit
<b>weights</b>	0	0.3027	0	0

Table 3: Minimum CVaR Portfolio

<b>Stock</b>	Atlantia	Autogrill	Mps	Enel
<b>weights</b>	0.0750	0	0	0
<b>Stock</b>	Eni	Fiat	Finmecc	Fondiaria
<b>weights</b>	0	0	0.0392	0
<b>Stock</b>	Generali	Intesa	Lottomat	Luxottica
<b>weights</b>	0	0	0	0.0705
<b>Stock</b>	Mediobanca	Mediolanum	Mediaset	Pirelli
<b>weights</b>	0.0884	0	0	0
<b>Stock</b>	Bpm	Saipem	Snam	Stm
<b>weights</b>	0	0	0.3913	0
<b>Stock</b>	Telecom	Terna	Ubi	Unicredit
<b>weights</b>	0	0.3356	0	0

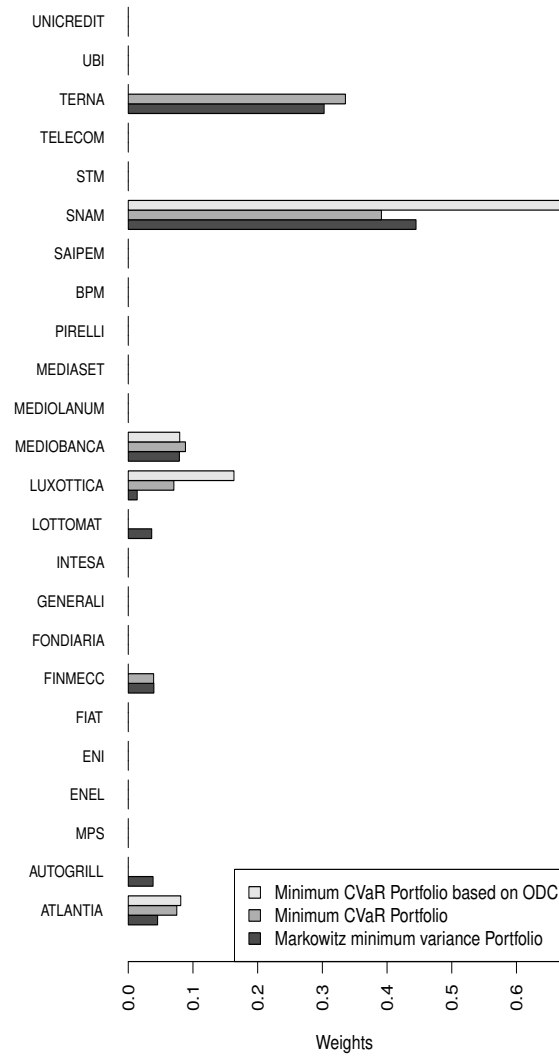


Figure 8: Weights of the stocks, Strategy A compared with classical portfolios

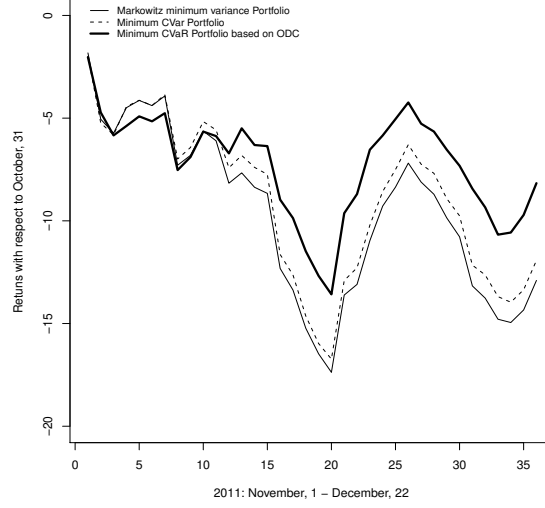


Figure 9: Portfolio returns, Strategy A compared with classical portfolios

7),  $T + 10$  (November 14),  $T + 15$  (November 21),  $T + 20$  (November 28),  $T + 25$  (December 5),  $T + 30$  (December 12),  $T + 35$  (December 19). At the end of day  $T + h$  we estimate the dissimilarity matrix for the following day,  $\Delta_{T+h+1}$ . In order to lighten the computational burden, the estimates of the tail dependence coefficients are computed using parameters estimated up to time  $T$  and price returns up to time  $T + i$ . In other words, we only update the information about the realized returns, and do not compute new estimates for the parameters of the involved statistical models. The dissimilarity matrix  $\Delta_{T+h+1}$  is then used to cluster the  $p$  stocks using the two-step clustering procedure described in Section 3.1. Let  $k$  be the number of clusters and  $n_1, \dots, n_k$  their sizes, we can build a total number of  $(n_1 \cdot \dots \cdot n_k)$  possible portfolios of  $k$  stocks using the criterion of

selecting stocks not belonging to the same cluster. For each of these  $(n_1 \dots n_k)$  possible selections, we estimate the weights of the Minimum CVaR Portfolios using returns at times  $1, \dots, T + h$  and we finally select the solution which exhibits the lowest CVaR value. The corresponding portfolio is then invested for the following 5 days, i.e. at times  $T + h + 1, \dots, T + h + 5$ . At the end of day  $T + h + 5$  the procedure is repeated and a new portfolio is built for investing during the following 5 days. The obtained portfolios are summarized in Table 4. In the short analysed period we do not observe appreciable changes in the composition of the portfolios, as the dynamic clustering in the out-of-sample data is fairly stable.

Also in this case, the returns of this investment strategy are compared to the returns of two competitors: a Markowitz minimum variance Portfolio and the Minimum CVaR Portfolio built using all the  $p = 24$  stocks and renewed every 5 days (Figure 10). The proposed strategy tends to outperform the other two.

Due to the substantial stability in the portfolios updated every 5 days, the returns deriving from the two proposed investment strategies, A and B, are very close each other (Figure 11). So, in this case study, the Minimum CVaR Portfolio based on ODC outperforms even the competitors (Markowitz minimum variance and Minimum CVaR) rebalanced every 5 days, which also suffer of higher transaction costs.

Table 4: Minimum CVaR Portfolios rebalanced every 5 days

Time: $T$	(October 31),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica	Snam	Atlantia
<b>weights</b>	0.0796	0.1634	0.6759	0.0811
Time: $T + 5$	(November 7),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica	Snam	Atlantia
<b>weights</b>	0.0398	0.1746	0.7100	0.0756
Time: $T + 10$	(November 14),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica	Snam	Atlantia
<b>weights</b>	0.0505	0.2009	0.7056	0.0430
Time: $T + 15$	(November 21),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica	Snam	Atlantia
<b>weights</b>	0.0380	0.1977	0.7082	0.0561
Time: $T + 20$	(November 28),	$k = 4$		
<b>Stock</b>	Ubi	Luxottica	Snam	Atlantia
<b>weights</b>	0	0.2187	0.7507	0.0306
Time: $T + 25$	(December 5),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica	Snam	Atlantia
<b>weights</b>	0.0417	0.1784	0.7039	0.0760
Time: $T + 30$	(December 12),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica	Snam	Atlantia
<b>weights</b>	0.0239	0.1972	0.7032	0.0757
Time: $T + 35$	(December 19),	$k = 4$		
<b>Stock</b>	Mediobanca	Luxottica 30	Snam	Atlantia
<b>weights</b>	0.0076	0.2093	0.7144	0.0687

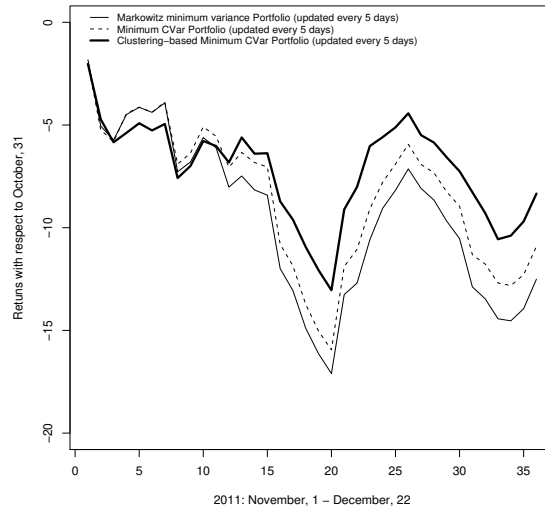


Figure 10: Portfolio returns, Strategy B compared with classical portfolios

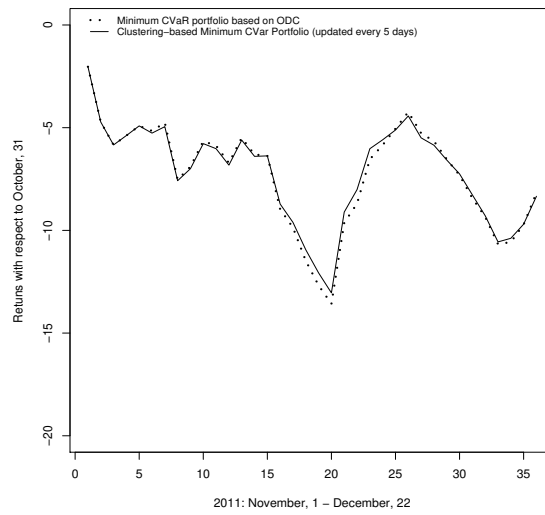


Figure 11: Portfolio returns, Strategy A and B



## 6 Discussion

In this paper we have proposed two portfolio selection strategies based on a dynamic clustering of the time series of returns observed for a set of candidate stocks. The dissimilarity matrix on which the stocks are clustered is derived from their lower tail dependence coefficients, estimated by means of copula functions with time-varying parameters. Thanks to this dissimilarity measure based on lower tail dependence, time series are clustered according to a similar behavior in event of extremely low returns, so that we propose, as a basic criterion, not to include in the portfolio stocks belonging to the same cluster. As a portfolio selection procedure, we propose to optimize the Conditional-Value-at-Risk, as this is consistent with the approach focused on extreme events.

The main advantages of the proposed method are:

- the time-varying copula functions model the relationship between lower tail dependence coefficients and the volatility of the market, so as to take into account contagion phenomenons;
- the diversification of the portfolio is strongly driven by the association of returns in the lower tail of their joint distribution, and this protects investments during financial crisis periods;
- the time-varying dissimilarity matrix results in a dynamic clustering solution which allows us a frequent portfolio rebalancement, if necessary.

A case study with real data from the Italian Stock Market during a financial crisis period shows that the two proposed procedures are able to outperform the

classical alternative methods. As future research, a criterion could be developed to decide which of the two strategies should be used in a given period and to automatically switch from one to another when it is the case. Furthermore, the procedure could be refined in order to take into account the upper tail dependence, which would allow us to take the best from bulls, while protecting from bears.

## References

- [1] Davies, D.L., and Bouldin, D. W. (1979), "A Cluster Separation Measure". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1, 224-227.
- [2] De Luca, G., and Zuccolotto, P. (2011), "A Tail Dependence-based Dissimilarity Measure for Financial Time Series Clustering". *Advances in Classification and Data Analysis*, 5, 323-340.
- [3] De Luca, G., and Zuccolotto, P. (2012), "Time Series Clustering on Lower Tail Dependence". *Proceedings of MAF2012, 2012: maf2012.unive.it/viewpaper.php?id=101*.
- [4] Dolničar, S., Leisch, F., Weingessel, A., Buchta, C., and Dimitriadou, E. A. (1998), "Comparison of Several Cluster Algorithms on Artificial Binary Data Scenarios from Tourism Marketing". *SFB Adaptive Information Systems and Modeling in Economics and Management Science*, Working Paper 7.

- [5] Dimitriadou, E., Dolničar, S., and Weingessel, A. (2002), "An Examination of Indexes for Determining the Number of Clusters in Binary Data Sets". *Psychometrika*, 67, 137-159.
- [6] Friedman, H. P., and Rubin, J. (1967), "On Some Invariant Criteria for Grouping Data". *Journal of the American Statistical Association*, 62, 1159-1178.
- [7] Hartigan, J. A., and Wong, M. A. (1979), "A K-means Clustering Algorithm". *Applied Statistics*, 28, 100-108.
- [8] Joe, H. *Multivariate Models and Dependence Concept*. Chapman & Hall, New York, 1997.
- [9] Krokmal, P., Palmquist, J., and Uryasev, S. (2002), "Portfolio Optimization with Conditional Value-at-Risk Objective and Constraints" *Journal of Risk*, 4, 43-68.
- [10] Nelsen RB. *An Introduction to Copulas*. Springer-Verlag, New York, 2006.
- [11] Rand WM. "Objective Criteria for the Evaluation of Clustering Methods". *Journal of the American Statistical Association*, 1971; 66:846-850.
- [12] Ratkowsky DA, Lance GN. "A Criterion for Determining the Number of Groups in a Classification". *Australian Computer Journal*; 1978; 10:115-117.
- [13] Rockafellar RT, Uryasev S. "Optimization of Conditional Value-at-Risk". *Journal of Risk*, 2000; 2:21-41.

- [14] Scott AJ, Symons MJ. "Clustering Methods Based on Likelihood Ratio Criteria". *Biometrics*, 1971; 27:387-397.
- [15] Sokal RR, Michener CD. "A Statistical Method for Evaluating Systematic Relationships". *University of Kansas Science Bulletin*, 1958; 38:1409-1438.